Surface water quality prediction system for Luton Hoo lake: *A statistical approach*

Tochukwu K. Anyachebelu[1], Marc Conrad[2] and Tahmina Ajmal[3]
Center for Wireless Research, Institute for Research in Applicable Computing
University of Bedfordshire
Luton, United Kingdom
[1]tochukwu.anyachebelu@study.beds.ac.uk, [2]marc.conrad@beds.ac.uk, [3]tahmina.ajmal@beds.ac.uk

Abstract—**Lake monitoring is a necessity for aquatic healthy living. Stakeholders are particularly interested not just in the aesthetics but also in the quality of water bodies. Our work tends to initially analyze historic data of the sensed water quality parameters at Luton Hoo lake to detect outliers. Dissolved oxygen has been predicted from available data since its one of the major surface water contaminants**

Keywords; Prediction, Sensors, correlations, water quality

I. INTRODUCTION

Water is one of the most vital natural resource in life [1]. The development of human civilization and evolution of life would have been impossible if not for water [2].

Achieving a good ecology, quantitative and qualitative status of surface waters has been made compulsory by the European Water framework Directive 2000/60(EU, 2000). Good water management plan should be able to show all significant pressures and the impact on aquatic environment by human activities [3].

Pollution affects the whole array of life in water and surface waters tend to be more vulnerable due to their easy accessibility for disposal of wastewaters. Industrialization, farming and domestic wastes are some of the anthropogenic activities that cause surface water deterioration [4].

The WFD has also identified the quality elements that requires monitoring in other to determine the surface water status. There are the biological elements, the hydro morphological elements and the physico-chemical elements which will be our major area of concern in this work [5]. The physico-chemical parameters to be monitored will be Temperature, Dissolved oxygen, Ammonia, conductivity, turbidity and pH level. Conditions necessary for the existence of a normal aquatic life has been stipulated by the WFD for various water bodies in different states [5].

A lot of research is being done using statistical analysis to assess the surface water quality and characteristics. This is seen as an easy and simple approach to monitoring all the water quality parameters [4].

A mechanism for prediction of these water quality parameters is provided through statistical regression analysis. This is a great tool for the correlation of different parameters [2]. It will economically aid the task of pollution status monitoring in surface water bodies, providing stakeholders with effective water quality management [2].

With the increasing volume of data from the sensors used for monitoring, there is need to use accurate data analysis. This will enable the development of faster and capable systems that will model and transform the huge data into essential knowledge for stakeholders [1].

Monitoring of the lakes remotely in real time has been made possible through the development of new sensor technologies and telemetry. Such installations has been carried out in some areas [6] to support environmental protection agencies in their decision making using real time information. The financial burden of realizing this real time monitoring in a secure, manageable and low cost manner is another issue to be considered [6].

Our present work is determined to study the water quality of the Luton Hoo Lake and ways of renovating it for the development of fisheries with other relevant activities for the stakeholders.

II. OVERVIEW

A. Description of study area

Luton Hoo Lake is a part of the River Lea. River Lea has one of the streams running from Leagrave into Hertfordshire and south into Greater London. Another stream runs from the west at Houghton Regis and continues south through Enfield and Tottenham Hale to Stratford [7]. The Luton Hoo Lake is divided into a north and south lake. It has two pollution sources which are the unmonitored run-off from sewage through Luton town and then the fuel etc. run-off from the Luton Airport. In February 2012 [8], a major pollution outbreak in one lake went un-noticed and resulted in the death of the entire fish stock. The lake cannot be restocked as no guarantee can be given that the same situation will not recur. The Environmental Agency has since installed 6-head sensor systems at the two points on the lake where pollution feeds into the lake. The Luton Hoo hotel management desires that the lake be renovated to add to the aesthetics of the estate.

The water quality parameters sensed by the installed sensors are Temperature, Conductivity, Dissolved oxygen, pH, Ammonium and Turbidity. Of great concern is the modeling of this lake to aid the prediction and early warning for contamination incidents. Our major stakeholders need to be

978-1-4799-4233-6/14/$31.00 ©2014 IEEE

alerted as early as possible so they can take actions to stop any type of contamination.

B. Water quality parameters and their dependencies

The various physico chemical parameters for determining water quality are discussed in this section:

Dissolved Oxygen (DO) in water is affected by many factors which include air, water temperature, wind mixing and photosynthetic activity. Other important factors to be considered are the oxidation process in water and sediments of the lake [9]. Salinity and temperature are natural water quality elements which affect the concentration of oxygen in the marine environment. The major cause of oxygen depletion is its consumption through the respiration of organisms living in the hypolimnion and oxidation of accumulated reduced substances, which usually occur over the summer. This directly affects the survival and growth of fishes in an aquatic environment [10]. The European Water Framework Directive classifications have resolved that water classification should be based on the concentration of dissolved oxygen (milligrams/Liter) instead of the percentage (%) saturation. This makes the outlined threshold to be dependent on the ambient temperature. Temperature increase as saturation increases while using percentage (%) saturation as a measure. Summer periods experience lower level of dissolved oxygen than any other season. The marine life environment is impacted with a great deal of stress at this period [11]. Threshold level is 6.0 milligrams/liter for warm water and 9.5 milligrams/Liter in cold water. Anything more than the threshold tends to be hazardous to the aquatic life [12].

Temperature of the lake is another factor to be considered. This controls the concentration of dissolved oxygen in the lake water [9]. An inverse relationship between the temperature value and dissolved oxygen level is deduced; where increase in temperature diminishes the solubility of oxygen [9]. The lake water temperature changes and its dynamics have a profound effect on the lake's biological and chemical processes [10]. More pollution incidents are expected in the summer months due to the high temperature within the period. 10oC increase in temperature at any time affects the life of fishes within the lake [12].

Ammonia in the lake exhibits highest concentrations during the winter period usually [13]. This is an important element to be monitored for water quality. Ammonium level detects the possibility of the presence of sewage, animal feces and bacterial pollution in surface water bodies. Human activities in areas around the lake also contribute to the ammonia content of surface water bodies. This may be as a result of agricultural, industrial and metabolic processes [14]. The current criteria for total ammonia concentrations are dependent on water temperature with pH between 6.5 and 9. The acute and chronic criterion are the basis for evaluation of the ammonium concentrations where the acute criterion is a 1-hour average concentration when salmonids are present and the chronic criterion is a 30 day average criterion when early life stage fishes are present [15]. Little dissolved oxygen in the lake adds to the toxicity of the ammonia content to the fishes. High pH and water temperature results in more toxic ammonia [16]. Ammonia level should remain greater than 0.1 milligram per liter and less than 1 milligram per liter for fishes to survive [12].

pH values recorded around June are relatively high. These are periods of warm water temperatures and fit with the algae bloom early onset. These periods experience high dissolved oxygen concentrations and high pH values as photosynthesizing algae consume carbon dioxide and produce oxygen [15]. Little vertical variation in pH values are noticed through the water column. This occurs at the surface most times when depth profiles are examined with the high pH [15]. Growing bloom results in supersaturated dissolved oxygen and high pH due to photosynthesizing algae consuming carbon dioxide and producing oxygen [15]. Higher pH values are caused by increased photosynthetic assimilation of dissolved inorganic carbon during summer. A similar effect is also caused by water evaporation. The use of detergents by people living around the lake for washing of cloths and utensils can be seen as source of more alkaline pH. Recommended pH level is from 6.5 – 9.0 [12].

Conductivity is the estimated total dissolved salts in water. It helps in deciding suitability of the water for irrigation and human consumption [11]. Conductivity is affected by inorganic dissolved solids like iron, calcium and organic compounds like oil, phenol which do not conduct electricity. It also depends on temperature whereby the warmer the water, the higher the conductivity. A sewage spill into the lake will raise the conductivity while an oil spill would lower the conductivity [12].

Turbidity influences the aesthetics of the lake. Soil erosion around the water banks cause suspended solids to enter the lake as well as inflow of effluent from sewage plants. Algae and plankton present in the water would also affect the turbidity levels. There tends to be an inverse relationship between turbidity and water depth depending on the area of a sampled region [17].

III. METHODS

This work is aimed at generating a model of the lake. Various researchers have developed similar models [18, 19, 20] for detecting trends in water quality parameters over a period of time. This would depend on the acquisition of water quality data and the application of appropriate statistical methods. The model would depend on thorough understanding of the relevant water quality relationships will form the basis of the model development.

The in-situ sensors are programmed to take readings of the various parameters at 15 minutes interval and send the data to the telemetry sink. The database stores this data for onward analysis and processing as the need arises. The base stations collate the readings from these monitoring equipment and sends them to the users through internet. Stakeholders can access this information on any desktop with internet connection by logging in with their details to view the real time chart of sensor readings. Users are able to extract previous data for analysis. Our subsequent work intends to create an alarm system based on the predicted values and real time value from monitored sites

Differentiating the trends from spurious noise using either parametric or nonparametric statistical method is of great importance in this work. The detected trends or extracted outliers from noise provides a basis for water contamination prediction. Trends refer to the permanent changes in the level of any particular water quality parameter while seasonal cycles are the oscillating patterns caused primarily by periodic changes in weather. Changes in natural biological activities and any managed activities influence concentrations of water quality patterns resulting in certain seasonal patterns. In evaluating changes in the water quality data, these seasonal effects should be considered and multiple regressions with periodic function would be a good way to describe the seasonal variations in parametric conditions.

Relationships between two continuous variables are established using the Correlation analysis. This is a systematic calculation for rapid water quality monitoring. The probability of linear relationship between X and Y is shown by the coefficient of correlation value near +1 and -1. Large values of correlation coefficient between two variables implies that they are highly correlated and this might be in the positive or negative direction [16]. Karl Pearson's correlation coefficient is used identify the highly correlated and inter related water quality parameters from the available water quality sensor data. To calculate the correlation coefficients for these parameters, correlation matrix was developed through the calculation of the coefficients for the different parameters measured. They are tested for significance using the p value where $p<0.05$, $p<0.01$ and non-significant when $p>0.05$. We considered the strong correlations when the value is greater than 0.5 and seen as a weak correlation when the value is less than 0.5.

Multiple linear regressions is a statistical tool for understanding the relationship between a dependent variable and one or more independent variables [21]. According to the researchers, multiple linear regressions can be expressed using the equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots + b_n X_n + \varepsilon \qquad (1)$$

Where Y represent the dependent variable;

$X_1, \cdots, X_n$ represent the several independent variables;

$\beta_0, \cdots, \beta_n$ represent the regression coefficient and;

ε represent the random error.

Understanding the problem clearly is important while choosing a statistical analysis. Descriptive statistics is an important aspect of analysis with regards to the central tendency of a set of data which includes the mean, median and mode. It is also necessary to identify the spread using the range and standard deviation [19]. Pearson correlation coefficients which considers normal distribution of the sensed water quality parameters is used in identifying the correlation structure between variables is studied using the Pearson correlation coefficient. Initial predictions are examined with the multiple linear regression. We discuss our main findings in the next section.

IV. RESULTS

The data acquired from installed sensors at the lake was examined and analyzed from October 2012 to March 2014. The correlations between various parameters determined for a 12 month period is shown in Figure 1.

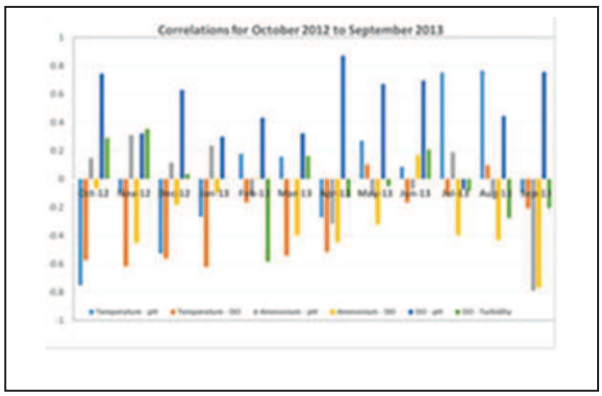


Figure 1 shows the correlation between different variables over one year - October 2012 to September 2013

There is a positive correlation between Temperature and Ammonium while negative correlations were noted for temperature with Dissolved oxygen, conductivity, pH and turbidity. Dissolved Oxygen has a positive correlation with conductivity, pH and turbidity while it exhibits a negative correlation with temperature and ammonium.

The negative correlation between Temperature and Dissolved oxygen is expected based on the high oxygen solubility in cold water. Strong negative correlations were noticed in October, November and December 2012, January, March, April and May 2013.

pH and dissolved oxygen is expected to give positive correlations as photosynthesis uses up the carbon dioxide in the water thereby reducing the pH level. Strong positive correlations were noticed in most of the months except in July and August 2013.

Temperature reading tends to move over the set points between April 2013 and October 2013 which are the peak of the sunny months in the UK. This remarkable change can be attributed to the weather change at the period which relatively affects the Dissolved oxygen level. Dissolved Oxygen was quite poor between June 2013 and September 2013 with the worst case happening in July and August 2013. Turbidity remained within an intermediate turbid level and getting to the medium turbid level in March 2013 and January 2014. Conductivity was within an acceptable level while Ammonium witnessed a spike in August 2013.

Ammonium and conductivity exhibited strong positive correlations for all the monitored period except May, June, July and August 2013 where the correlations were not strong. This lead to the closer examination of August 2013 to detect

possible anomalies as seen in the correlation and telemetry chart in Figure 2 and 3 respectively. The monthly chart was split into four sections of the day on a 6-hourly basis to help give a closer highlight of outliers. They are Morning (6am – 11.59am), Afternoon (12pm – 5.59pm), Evening (6pm – 11.59pm) and Night (12am – 5.59am).

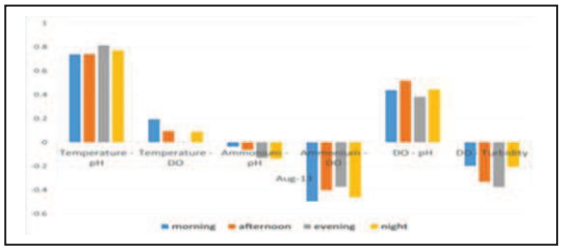


Figure 2: August 2013 Correlation data with contamination threats

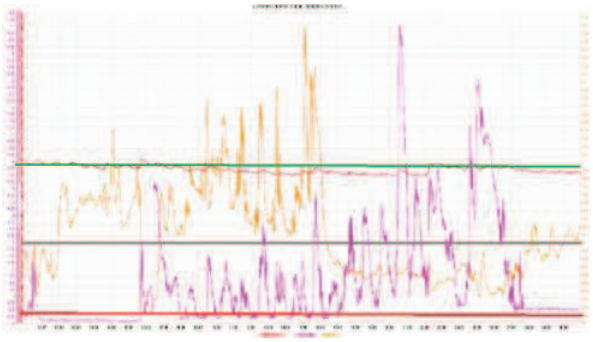


Figure 3: Screenshot of the Water monitoring telemetry (August 2013)– Red line marks the lower set point for Ammonia while Green marks the lower set point for Dissolved oxygen and Grey is the upper limit for Ammonia

This is further compared with October 2012 correlation and telemetry chart in figure 5 and 6 respectively. There was stronger correlations for the parameters sensed.

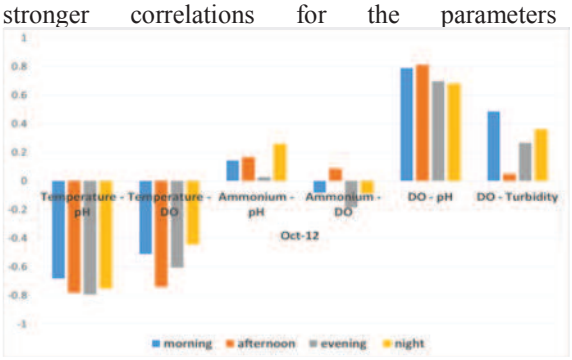


Figure 4: October 2012 Correlation data at normal state



Figure 5: Screenshot of the Water monitoring telemetry (Oct 2012) – Red line marks the lower set point for Ammonium while Green marks the lower set point for Dissolved oxygen

In the multiple linear regression analysis, Dissolved oxygen is taken as the dependent variable since it is a major pollution dictator while other parameters are examined as the independent variables. The R square (R2) as derived from the analysis gives the percentage of variability in the dependent variable accounted to all the independent variables as 72.1%. This is the measure of the strength of association with all parameters examined as shown in the Table 1. The F-test shows if the model is a good fit for the data with the p-value showing the overall significance compared to the alpha level of 0.05.

The independent variables are taken as the predictors which gives the individual unstandardized coefficients for each. In this analysis, the standardization of coefficients determines which of the variables has a greater effect on the dependent variable (Dissolved oxygen) when measured in their different units. The standardized coefficients of a relationship represent its strength relative to the variation in the distributions. Unstandardized coefficients are used in deriving the equation for the regression as all the variables cannot be measured in same units as in Table 2.

The regression equation for the entire 18 months (October 2012 to March 2014) period is given as (2):

Dissolved oxygen = -45.819 -0.443Temperature + 0Conductivity +7.814pH -0.855Ammonia +0.002Turbidity + 1.755036 + ε (2)

For every unit increase in temperature, -0.443 unit decrease in dissolved oxygen is predicted where all other variables are constant. Also for every unit of increase in pH, a 7.814 unit increase in dissolved oxygen is predicted with all other variables constant. There is a -0.855 unit decrease in dissolved oxygen for every unit increase of ammonia. Unit increase of conductivity does not affect the dissolved oxygen predicted value where other variables are held constant.

From the equation, it is noticed that pH and turbidity have a positive effect on the predicted dissolved oxygen while temperature and ammonia have a negative effect which gives an indication of various sources of pollution. All the parameters exhibit a significant level based on the p-values which are less than 0.05. However the various months are examined to determine irregularities for different periods which are shown in the Table 3.

CONCLUSION AND DISCUSSION

A study of the main water quality parameters for a one year duration and examined correlations between them. It was observed that some of the correlations undergo seasonal changes with temperature, dissolved oxygen and ammonia going beyond the stipulated limits especially during the summer.

We have made a comparison for two months. There are sharp peaks which are seen as result of contamination thereby giving an indication of a trend that could lead to pollution prediction. In analyzing the sensor data in this work, we have been able to see the trends and outliers which are to be monitored for new data. The correlation analysis has been able to show definite parameters that needs to be frequently measured and considered in the model building for efficient determination of the water quality status.

For this work we have used the telemetry data, however using this in real time context will facilitate further validation of the predicted values for the water quality parameters. This provides input to decision support systems with regards to abnormalities leading to the surface water contamination.

The performance of the modelled lake is likely to exhibit great efficiency with multiple sensors installed at different points of the lake and the collation of sensed variables for decision making. It essentially gives a greater advantage than having to take water samples to the laboratory for analysis whereas you can get an on-the-spot result for effective management.

Dissolved oxygen value has been predicted using the multiple linear regression where all other parameters are seen as the dependent variable for the prediction.

From this study we conclude that a hybrid model of neural network and autoregressive integrated moving average (ARIMA) could be used in the further prediction of the water quality parameters owing to its Time series factor [22]. This will give room to further validation of predicted values to ensure the correctness of prediction. Neural network is considered suitable for classification, forecasting, prediction and is highly accepted due to its ability to treat complicated and nonlinear problems which are difficult to evaluate using the traditional water quality measurement models [23]. Defining the number of hidden neurons for a particular problem is the most challenging part and MATLAB neural network toolbox gives a platform for this determination. ARIMA will deal with both the seasonal and non-seasonal part of the model on the assumption that further values of the water quality parameters are linear functions of several past observations.

REFERENCES

[1] P. Dutta and R. Chaki, "A survey of data mining applications in water quality management," in Proceedings of the CUBE International Information Technology Conference, 2012, pp. 470-475.

[2] N. Kumar and D. Sinha, "An Approach to River Water Quality Management through Correlation Study among Various Water Quality Parameters," Int J Environ Sci, vol. 1, pp. 253-259, 2010.

[3] M. Diamantopoulou, V. Antonopoulos and D. Papamichail, "The use of a neural network technique for the prediction of water quality parameters of Axios river in northern Greece," European Water, vol. 11, pp. 55-62, 2005.

[4] N. Khatoon, A. H. Khan and M. R. Pathak, "Correlation Study For the Assessment of Water Quality and Its Parameters of Ganga River, Kanpur, Uttar Pradesh, India,"

[5] M. Best, A. Wither and S. Coates, "Dissolved oxygen as a physico-chemical supporting element in the Water Framework Directive," Mar. Pollut. Bull., vol. 55, pp. 53-64, 2007.

[6] X. Zhu, D. Li, D. He, J. Wang, D. Ma and F. Li, "A remote wireless system for water quality online monitoring in intensive fish culture," Comput. Electron. Agric., vol. 71, pp. S3-S9, 2010.

[7] River lea Accessed online at http://river-lea.co.uk/

[8] BBC News of 08/03/2012 Hundreds of fish at Luton Hoo lakes killed by sewage [online] Accessed at: http://www.bbc.co.uk/news/uk-england-beds-bucks-herts-17297136

[9] A. Al Sayes, A. Radwan and L. Shakweer, "Impact of drainage water inflow on the environmental conditions and fishery resources of Lake Borollus," 2007.

[10] V. Z. Antonopoulos and S. K. Gianniou, "Simulation of water temperature and dissolved oxygen distribution in Lake Vegoritis, Greece," Ecol. Model., vol. 160, pp. 39-53, 2003.

[11] P. J. Puri, M. Yenkie, D. Battalwar, N. V. Gandhare and D. B. Dhanorkar, "Study and Interpretation of Physico-Chemical Characteristic of Lake Water Quality in Nagpur City (India)," Rasayan J.Chemistry, vol. 3, pp. 800-810, 2010.

[12] Parliament. UK Technical Advisory Group on the Water Framework Directive. (2013) Updated R Ecommendations on Environmental Standards River Basin Management (2015 - 21). UNITED KINGDOM: WFD UK TAG.

[13] B. Alejandra, L. Shuming and V. Francois, "Drinking Water Source Contamination Early Warning System and Modelling in China: A Review," International Journal of Environmental Pollution and Remediation, vol. 1, pp. 13-19, 2012.

[14] Q. Fu, B. Zheng, X. Zhao, L. Wang and C. Liu, "Ammonia pollution characteristics of centralized drinking water sources in China," Journal of Environmental Sciences, vol. 24, pp. 1739-1743, 2012.

[15] S. Hostetler, "Use of models and observations to assess trends in the 1950–2005 water balance and climate of Upper Klamath Lake, Oregon," Water Resour. Res., vol. 45, 2009.

[16] A. Bhatnagar and P. Devi, "Applications of correlation and regression analysis in assessing lentic water quality: a case study at Brahmsarovar Kurukshetra, India." International Journal of Environmental Sciences, vol. 3, 2012.

[17] G. L. Howick and J. Wilhm, "Turbidity in lake carl blackwell: Effects of water depth and wind," in Proc. Okla. Acad. Sci, 1985, pp. 51-57.

[18] D. Hou, X. Song, G. Zhang, H. Zhang and H. Loaiciga, "An early warning and control system for urban, drinking water quality protection: China's experience," Environmental Science and Pollution Research, vol. 20, pp. 4496-4508, 2013.

[19] The lake and Reservoir Restoration Guidance manual 1993 "Statistical methods for the analysis of lake water quality trends" Accessed online at http://nepis.epa.gov/Exe/ZyPDF.cgi/20004RZ7.PDF?Dockey=20004RZ7.PDF on 30/04/2014

[20] L. Fu and Y. Wang, "Statistical Tools for Analyzing Water Quality Data," .

[21] S. Basu and K. Lokesh, "Application of Multiple Linear Regression and Manova to Evaluate Health Impacts Due to Changing River Water Quality," Applied Mathematics, vol. 5, pp. 499, 2014.

[22] D. Ömer Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," Eng Appl Artif Intell, vol. 23, pp. 586-594, 2010.

[23] Li Ying, Zhou Jiti, Wang Xiangrui and Zhou Xiaohui, "Water quality evaluation of nearshore area using artificial neural network model," in Bioinformatics and Biomedical Engineering , 2009. ICBBE 2009. 3rd International Conference on, 2009, pp. 1-4.

Table 1 showing the R2 and standard error for the period investigated

Model	R	R square	Adjusted R square	Standard error of the estimate
1	0.849	0.721	0.721	1.755036

Table 2 showing Unstandardized Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for β	
		β	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-45.819	.374		-122.642	.000	-46.551	-45.087
	temperature	-.443	.003	-.404	-126.760	.000	-.450	-.436
	conductivity	.000	.000	-.026	-9.649	.000	.000	.000
	pH	7.814	.047	.510	166.415	.000	7.722	7.906
	Ammonium	-.855	.032	-.074	-26.683	.000	-.918	-.793
	turbidity	.002	.000	.022	9.295	.000	.001	.002

Table 3 showing the derived regression equation for individual months.

Month	Regression equation	R^2	F value
October 2012	-33.309 – 0.304temp – 0.005Conductivity + 6.236pH +0.522Ammonia – 0.001Turbidity + 0.699998	0.773	1432.791
November 2012	-29.147 – 0.597temp – 0.002Conductivity + 5.908pH – 2.419Ammonia – 0.001Turbidity + 0.878931	0.616	916.462
December 2012	-14.578 – 0.283temp + 0Conductivity + 3.645pH - 1.661Ammonia – 0.004Turbidity + 0.489273	0.681	1263.156
January 2013	-22.018 – 0.353temp – 0.002Conductivity + 4.754pH + 1.355Ammonia – 0.001Turbidity + 0.527440	0.602	896.154
February 2013	-10.639 – 0.461temp – 0.001Conductivity + 3.258pH +0.241Ammonia – 0.007Turbidity + 0.542770	0.544	637.781
March 2013	-33.431 – 0.454temp + 0.001Conductivity + 6.260pH – 2.114Ammonia + 0.003Turbidity + 0.783971	0.517	633.987
April 2013	-131.609 – 0.153temp + 0.001Conductivity + 18.896pH – 2.596Ammonia + 0.009Turbidity + 0.906531	0.888	4557.622
May 2013	-93.851 + 0.367temp – 0.005Conductivity + 13.684pH – 2.217Ammonia – 0.002Turbidity + 1.231057	0.699	1372.442
June 2013	-58.397 – 0.006temp – 0.004Conductivity + 8.020pH +1.898Ammonia – 0.001Turbidity + 1.102295	0.509	589.045
July 2013	-11.655 + 0.495temp + 0.007Conductivity + 0.346pH -3.482Ammonia + 0.032Turbidity + 1.612651	0.357	329.047
August 2013	-46.097 – 1.039temp – 0.005Conductivity + 9.676pH -0.836Ammonia – 0.008Turbidity + 1.253012	0.584	812.074
September 2013	-32.561 – 0.361temp – 0.006Conductivity + 6.076pH +0.331Ammonia + 0.003Turbidity + 1.014070	0.797	2238.596
October 2013	-58.969 – 0.196temp – 0.004Conductivity + 9.290pH – 3.754Ammonia – 0.002Turbidity + 0.802256	0.870	3945.125
November 2013	-56.513 – 0.565temp – 0.004Conductivity + 9.381pH +1.634Ammonia – 0.020Turbidity + 0.951566	0.620	921.469
December 2013	-36.273 – 0.338temp – 0.003Conductivity + 6.668pH – 4.894Ammonia – 0.001Turbidity + 0.856798	0.810	2485.670
January 2014	-12.701 – 0.50temp – 0.001Conductivity + 3.671pH – 0.860Ammonia – 0.001Turbidity + 0.501502	0.597	878.383
February 2014	-13.058 – 0.344temp – 0.002Conductivity + 3.639pH – 0.364Ammonia + 0Turbidity + 0.295353	0.769	1779.609
March 2014	-36.173 – 0.374temp – 0.004Conductivity + 6.816pH +2.732Ammonia + 0Turbidity + 0.271035	0.755	621.440